# Zero Trust &
# Zero Trust Architecture

ICT Risk Assessment
04/07/2022

Leonardo Vona
545042

# Zero Trust Approach

- Zero trust is not a single architecture but a set of guiding principles
- Focused on assets and subjects protection
- From an operative point of view, “*zero trust provides a collection of concepts and ideas designed to minimize uncertainty in enforcing accurate,* ***least privilege, per-request*** *access decisions in information systems and services in the face of a network viewed as compromised*”
- Assume that
  - an attacker is present in the environment
  - There is no implicit trust: an organization-owned environment is no more trustworthy than any nonorganization-owned environment
- The risks have to be continually analyzed and evaluated
- Usually, access to resources is minimized to only those subjects and assets identified as needing access as well as continually authenticating and authorizing the identity and security posture of each access request

# Zero Trust Architecture

- A zero trust architecture is *“an enterprise cybersecurity architecture that is based on zero trust principles and designed to prevent data breaches and limit internal lateral movement”*
- The initial effort is on restricting resources to those with a need to access and grant only the **minimum** privileges needed to perform the mission
  - Traditionally, perimeter defense and authenticated subjects are given authorized access to a broad collection of resource once on the internal network. This leads to unauthorized lateral movement within the environment
- The goal is to prevent unauthorized access to resources and services coupled with making the access control enforcement **as granular as possible**
  - The focus is on authentication, authorization, and shrinking implicit trust zones while maintaining availability and minimizing temporal delays in authentication mechanisms. Access rules are made as granular as possible to enforce least privileges needed to perform the action in the request.

# Tenets of Zero Trust

- Many definitions of zero trust stress the concept of removing wide-area perimeter defenses as a factor. The following is an attempt to define ZT and ZTA in terms of basic tenets that should be involved rather than what is excluded

  ***1) All data sources and computing services are considered resources***

  ***2) All communication is secured regardless of network location.*** Network location alone does not imply trust.

  ***3) Access to individual resources is granted on a per-session basis.*** Trust in the requester is evaluated before the access is granted. Access should also be granted with the least privileges needed to complete the task. Authentication and authorization to one resource will not automatically grant access to a different resource

# Tenets of Zero Trust (cont.)

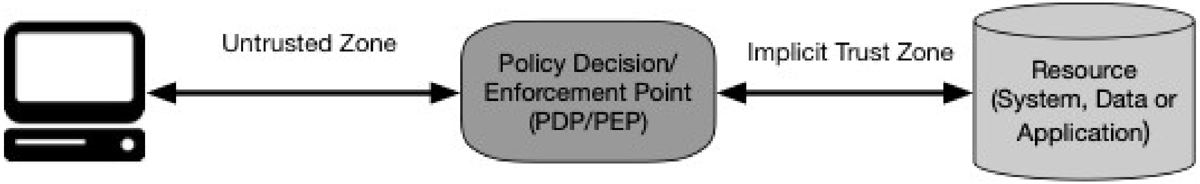
4) ***Access to resources is determined by dynamic policy, and may include other behavioral and environmental attributes.*** *Least privilege principles are applied to restrict both visibility and accessibility*

5) ***The organization monitors and measures the integrity and security posture of all owned and associated assets****. No asset is inherently trusted. The organization evaluates the security posture of the asset when evaluating a resource request*

6) ***All resource authentication and authorization are dynamic and strictly enforced before access is allowed***

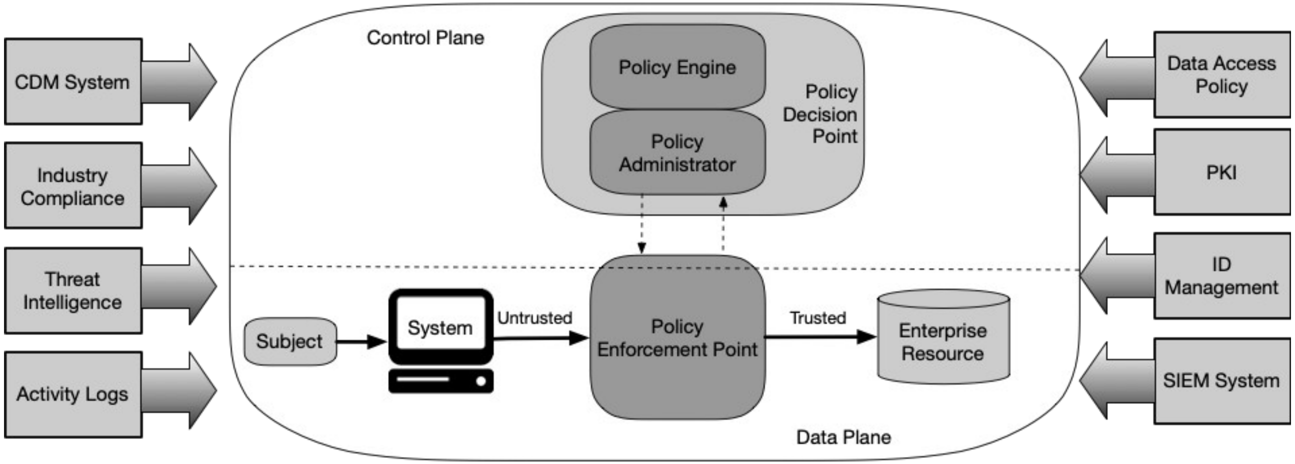
7) ***The organization collects as much information as possible about the current state of assets, network infrastructure and communications, and uses it to improve its security posture***

# Access Abstract Model


Untrusted Zone
Policy Decision/ Enforcement Point (PDP/PEP)
Implicit Trust Zone
Resource (System, Data or Application)

- The access to a resource is granted through a policy decision point (PDP) and corresponding policy enforcement point (PEP).
- The system must ensure that the subject requesting the access is authentic and the request is valid. The PDP/PEP passes proper judgement to allow the subject to access the resource. This implies that zero trust applies to **authentication** and **authorization**
- Organizations need to develop and maintain **dynamic risk-based policies** for resource access and set up a system to ensure that these policies are enforced correctly and consistently for **individual** resource access requests
- Zero trust provides a set of principles and concepts around moving the PDP/PEPs closer to the resource. The idea is to **explicitly authenticate and authorize all subjects**, assets and workflows that make up the organization
  - The “*implicit trust zone*” represents an area where all entities are trusted to at least the level of the last PDP/PEP gateway
  - The PDP/PEP applies a set of controls so that all traffic beyond PEP has a common level of trust. To allow the PDP/PEP to be as specific as possible, *the implicit trust zone must be as small as possible*

# Logical Components of ZTA



- There are numerous logical components that make up a ZTA deployment in an organization. These components may be operated as an on-premises services or through a cloud-based services

# Logical Components of ZTA (cont.)

- **Policy decision point**
  - ***Policy engine (PE)***: component responsible for the decision to grant access to a resource for a given subject. The PE uses organization policy as well as data from external sources as input to a *trust algorithm*. The PE makes and logs the decision, and the PA executes the decision
  - ***Policy administrator (PA)***: component responsible for establishing and/or shutting down the communication path between a subject and a resource (via commands to relevant PEPs). It would generate any authentication token or credential used by a client to access an organization resource. The PA communicates with the PEP when creating the communication path
- **Policy enforcement point**: system responsible for enabling, monitoring, and eventually terminating connections between a subject and an organization resource. The PEP communicates with the PA to forward requests and/or receive policy updates from the PA. This is a single logical component in ZTA but may be broken into two different components: the client (e.g., agent on a laptop) and resource side (e.g., gateway component in front of resource that controls access). Alternatively is can be a single portal component that acts as a gatekeeper for communication paths. Beyond the PEP is the trust zone.

# Variations of ZTA Approaches

- There are several ways an organization can enact a ZTA for workflows. These approaches vary in the components used and in the main source of policy rules for an organization. A full ZT solution will include elements of all three approaches
  - ***ZTA using enhanced identity governance***: the primary requirement for resource access is based on the access privileges granted to the given subject. Organization resource access policies are based on identity and assigned attributes.
    - This approach is often deployed using an open network model: network access is initially granted to all assets but access to organization resources are restricted to identities with the appropriate access privileges
    - The downside in granting basic network connectivity is that malicious actors could still attempt network reconnaissance and/or use the network to launch DoS attacks
    - Identity-driven approaches also work well for organizations that use cloud-based services that may not allow for organization-owned or -operated ZT security components to be used

# Variations of ZTA Approaches (cont.)

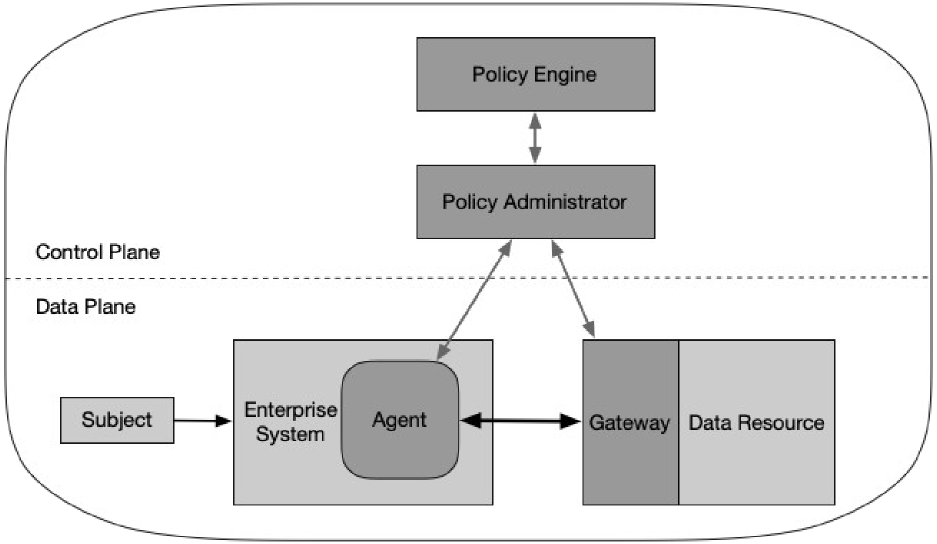
- ***ZTA using micro-segmentation***: based on placing individual or groups of resources on a unique network segment protected by a gateway security component. In this approach, the organization places infrastructure devices to act as PEPs protecting each resource or small group of related resources
  - This approach relies on the gateway components to act as the PEP that shields resources from unauthorized access and/or discovery
- ***ZTA using network infrastructure and software defined perimeters***: the ZTA implementation could be achieved by using an overlay network. In this approach, the PA acts as the network controller that sets up and reconfigures the network based on the decisions made by the PE. The clients continue to request access via PEPs
  - When the approach is implemented at the application network layer, the most common deployment model is the agent/gateway

# Deployment Variations

- The above components are logical; they do not necessarily need to be unique systems. There are several variations on deployment of selected components of the architecture. Some of them are:
  - Device agent/gateway-based deployment
  - Enclave-based deployment
  - Resource portal-based deployment
  - Device application sandboxing

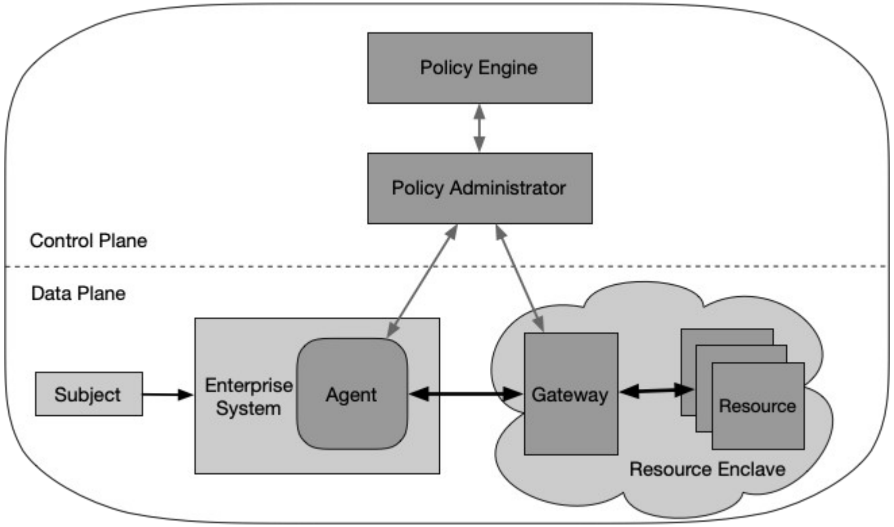
# Device Agent/Gateway-Based

- The PEP is divided into two components: one residing on the asset and one directly in front of a resource. For example, each organization-issued asset has an installed device agent that coordinates connections, and each resource has a component (the gateway) that is placed directly in front, so that the resource communicates only with the gateway, essentially serving as a proxy for the resource.
- This model is best utilized for organizations that have a robust device management program in place



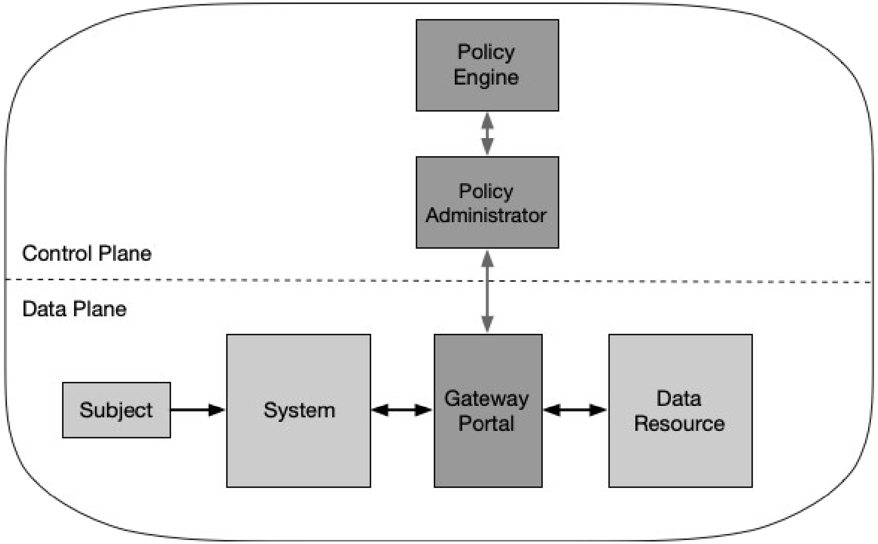
# Enclave-Based

- Is a variation of the device agent/gateway model. The gateway components may not reside on assets or in front of individual resources but instead reside at the boundary of a resource enclave. In this model, the entire private cloud is located behind a gateway
- This model is useful for organizations that have legacy applications or on-premises data centers that cannot have individual gateways in place
- The downside is that the gateway protects a collection of resources and may not be able to protect each resource individually. This may also allow for subjects to see the resources which they do not have privileges to access



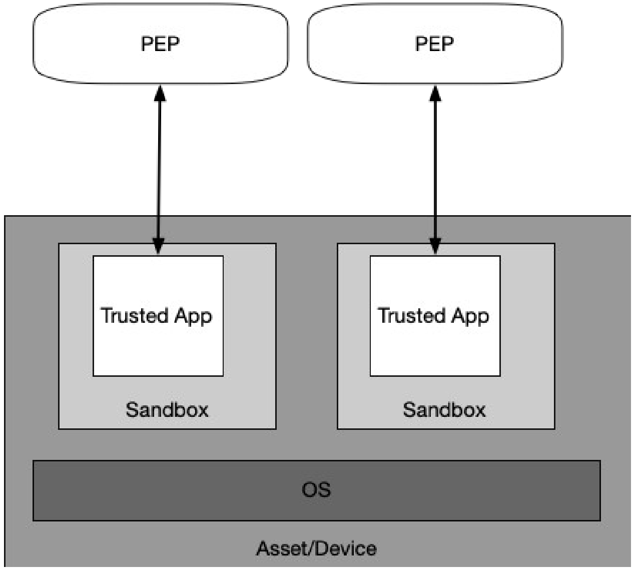
# Resource Portal-Based

- The PEP is a single component that acts as a gateway for subject requests
- The primary benefit of this model is that a software component does not need to be installed on all client devices
- However, limited information can be inferred from devices requesting access. This model may not be able to continuously monitor assets for malware, unpatched vulnerabilities, and appropriate configuration
- This model also allows attackers to discover and attempt to access the portal or attempt a DoS attack against the portal. The portal system should be well-provisioned



# Device Application Sandboxing

- Another variation of the agent/gateway model, having vetted applications or processes run compartmentalized on assets.
- The goal is to protect the application from a possibly compromised host
- The device runs approved, vetted applications in a sandbox. The applications can communicate with the PEP to request access to resources, but the PEP will refuse requests from other applications on the asset
- The main advantage of this model is that individual applications are segmented from the rest of the asset
- One of the disadvantages of this model is that organizations must maintain these sandboxed applications for all assets and may not have full visibility into client assets. The organization also needs to make sure each sandboxed application is secure, which may require more effort than simply monitoring devices



# Trust Algorithm

- The PE’s trust algorithm takes input from multiple sources, which can be grouped into broad categories:
  - ***Access request***: actual request from the subject
  - ***Subject database and history***: set of subjects of the organization and relative attributes/privileges assigned
  - ***Asset database (and observable status)***: contains the known status of each organization-owned (and possibly non-organization/BYOD) asset
  - ***Resource policy requirements***: set of policies which complements the user ID and attributes database and defines the minimal requirements for access the resource
  - ***Threat intelligence and logs***: feed(s) about general threats and active malware operating on the Internet
- It is important to balance security, usability, and cost-effectiveness when defining and implementing trust algorithms

# Trust Algorithm (cont.)

- There are two characteristics that can be used to differentiate TAs: how the factors are evaluated and how the request is evaluated in relation on other requests by the same subject
  - ***Criteria- versus score-based***
    - A criteria-based TA assumes a set of qualified attributes that must be met before access is granted to a resource
    - A score-based TA computes a confidence level based on values for every data source and organization-configured weights. It the score is greater that the configured threshold value for the resource, access is granted
    - Developing a set of criteria or weights/threshold values for each resource requires planning and testing
  - ***Singular versus contextual***:
    - A singular TA treats each request individually. This can allow faster evaluations, but there is a risk that an attack can go undetected if it stays within a subject’s allowed role
    - A contextual TA takes the subject recent history into consideration. This means the PE must maintain some state information on all subjects but may be more likely to detect an attacker accessing information in a pattern that is atypical of what the PE sees for the given subject. This also means that the PE must be informed of user behavior by the PA (and PEPs)
- Contextual, score-based TAs would provide the ability to offer more dynamic and granular access control

# Migrating to a ZTA

- An organization should seek to incrementally implement ZT principles. Most organizations will continue to operate in a hybrid ZT/perimeter-based mode for an indefinite period
  - Migration to a ZTA approach to the organization may take place one business process at a time
  - The organization needs to make sure that the common elements are flexible enough to operate in a ZTA and perimeter-based hybrid security architecture
- An organization should reach a baseline of competence before it becomes possible to deploy a significant ZT-focused environment. This baseline includes having assets, subjects, business processes, traffic flows and dependency mappings identified and cataloged for the organization.
- Incomplete knowledge will most often lead to a business process failure where the PE denies requests due to insufficient information

# Migrating to ZTA (cont.)

- Steps for the migration to ZTA:
  - ***Identify actors on the organization***: the PE must have knowledge of organization subjects
  - ***Identify assets owned by the organization***: one of the key requirements of ZTA is the ability to identify and manage devices. The ability to observe the current state of an asset is part of the process of evaluating access requests
  - ***Identify key processes and evaluate risks associated with executing process***: consider potential trade offs in performance, user experience, and possible increased workflow fragility that may occur when implementing ZTA for a given business process
  - ***Formulating policies for the ZTA candidate***: after the asset or workflow is identified, identify all upstream resources (e.g. ID management systems, databases, microservices), downstream resources (e.g., logging, security monitoring), and entities (e.g., subjects, service accounts) that are used or affected by the workflow. The organization administrators then need to determine the set of criteria or confidence level weights for the resources used in the candidate business process

# Migrating to ZTA (cont.)

- ***Identify candidate solutions***: there are some factor to consider, such as
  - Does the solution require that components be installed on the client asset?
  - Does the solution require changes to subject behavior?
- ***Initial deployment and monitoring***: organization administrators must implement the developed policies by using the selected components but may wish to operate in an observation and monitoring mode at first
- ***Expanding the ZTA***: the network and assets are still monitored, and traffic is logged, but responses and policy modifications are done at a lower tempo as they should not be severe. At this stage, the organization administrators can begin planning the next phase of ZT deployment. If a change occurs to the workflow, the operating ZT architecture needs to be reevaluated.

# Considerations about ZTA

- ZTA planning and implementation may change the authorization boundaries defined by the organization. This is due to the addition of new components (e.g., PE, PA, and PEPs) and a reduction of reliance on network perimeter defenses
- Part of the core requirements for ZT is that an organization should inspect and log traffic in its environment. Some of this traffic may contain private information or have associated privacy risks
- Strong subject provision and authentication policies need to be in place before moving to a more ZT-aligned deployment. Organizations need a clear set of subject attributes and policies that can be used by a PE to evaluate access requests
- Having a strong CDM (Continuous Diagnostics and Mitigation) program implementation is key to the success of ZTA. For example, to move to ZTA, an organization must have a system to discover and record physical and virtual assets to create a usable inventory

# Threats associated with ZTA

- ZTA can reduce overall risk and protect against common threats. However, some threats have unique features when implementing a ZTA
  - ***Subversion of ZTA decision process***: Any organization administrator with configuration access to the PE’s rules may be able to perform unapproved changes or make mistakes that can disrupt organization operations. Likewise, a compromised PA could allow access to resources that would otherwise not be approved. The PE and PA components must be properly configured and monitored
  - ***DoS or network disruption***: If an attacker disrupts or denies access to the PEP(s) or PE/PA, it can adversely impact organization operations. Organizations can mitigate this threat by having the policy enforcement reside in a properly secured cloud environment or be replicated in several locations. There is also the risk that organization resources may not be reachable from the PA
  - ***Stolen credentials/Insider threat***: Accounts with access policies around resources that an attacker is interested in would be the primary targets for attackers. Implementation of MFA for access requests may reduce the risk of information loss from a compromised account. ZTA reduces risk and prevents any compromised accounts or assets from moving laterally throughout the network

# Threats associated with ZTA (cont.)

- ***Visibility on the network***: The organization may not perform deep packet inspection or examine the encrypted traffic and must use other methods to assess a possible attacker on the network
- ***Storage of system and network information***: A related threat to organization monitoring and analysis of network traffic is the analysis component itself. If monitor scans, network traffic, and metadata are being stored, that data becomes a target for attackers If an attacker gains access to this information, it may be able to identify assets for further reconnaissance and attack
- ***Reliance on proprietary data formats or solutions***: Often, the assets used to store and process this information do not have a common, open standard on how to interact and exchange information. This can lead to instances where an organization is locked into a subset of providers due to interoperability issues
- ***Use of non-person entities (NPE) in ZTA administration***: Artificial intelligence and other software-based agents are being deployed to manage security issues on organization networks. How these components authenticate themselves in an organization implementing a ZTA is an open issue. The associated risk is that an attacker will be able to induce or coerce an NPE to perform some task that the attacker is not privileged to perform